

Das ist der HAMMER: Consistent new physics interpretations of semileptonic decays

Florian U. Bernlochner^{a,1}, Stephan Duell^{b,1}, Zoltan Ligeti^{c,2},
Michele Papucci^{d,2,3}, Dean J. Robinson^{e,2}



¹Physikalisches Institut der Rheinischen Friedrich-Wilhelms-Universität Bonn, 53115 Bonn, Germany

²Ernest Orlando Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA

³Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, USA

Abstract Precise measurements of $b \rightarrow c\tau\bar{\nu}$ decays require large resource-intensive Monte Carlo (MC) samples, which incorporate detailed simulations of detector responses and physics backgrounds. Extracted parameters may be highly sensitive to the underlying theoretical models used in the MC generation. Because new physics (NP) can alter decay distributions and acceptances, the standard practice of fitting NP Wilson coefficients to SM-based measurements of the $R(D^{(*)})$ ratios can be biased. The newly developed **Hammer** software tool enables efficient reweighting of MC samples to arbitrary NP scenarios or to any hadronic matrix elements. We demonstrate how **Hammer** allows avoidance of biases through self-consistent fits directly to the NP Wilson coefficients. We also present example analyses that demonstrate the sizeable biases that can otherwise occur from naive NP interpretations of SM-based measurements. The **Hammer** library is presently interfaced with several existing experimental analysis frameworks and we provide an overview of its structure.

Contents

1	Introduction	1
2	New physics analyses	2
2.1	MC sample	3
2.2	Reweighting and fitting analysis	3
2.3	$R(D^{(*)})$ biases from new physics truth	4
2.4	New physics Wilson coefficient fits	5
3	The Hammer library	6
3.1	Reweighting	7
3.2	New Physics generalizations	7
3.3	Form factor generalizations	8

3.4	Rates	9
3.5	Primary code functionalities	9
3.6	Code flow	10
4	Conclusions	10
A	Core elements of the Application Programming Interface	11
A.1	Building processes and events	11
A.2	Specifications	12
A.3	Histogramming	13
A.4	Processing	14
A.5	Setting Wilson coefficients and form factors	14
A.6	Retrieval	14
A.7	Multithreading	15
A.8	Saving	15
A.9	Reloading and merging	16

1 Introduction

Precision analyses of semileptonic b -hadron decays typically rely on detailed numerical Monte Carlo (MC) simulations of detector responses and acceptances. Combined with the underlying theoretical models, these simulations provide MC *templates* that may be used in fits, to translate experimental yields into theoretically well-defined parameters. This translation though can become sensitive to the template and its underlying theoretical model, introducing biases whenever there is a mismatch between the theoretical assumptions used to measure a parameter and subsequent theoretical interpretations of the data.

Such biases are known to arise in the analyses of semileptonic decays of b hadrons, in particular, for the measurements of the CKM element $|V_{cb}|$, and the ratio of semitaonic vs. semileptonic decays to light leptons,

$$R(H_c) = \frac{\Gamma(H_b \rightarrow H_c \tau \bar{\nu})}{\Gamma(H_b \rightarrow H_c l \bar{\nu})}, \quad l = \mu, e, \quad (1)$$

where $H_{b,c}$ denote b - and c -flavor hadrons. To avoid this, the size of these biases need to be either carefully controlled when experiments quote their results

^aflorian.bernlochner@uni-bonn.de

^bs.duell@physik.uni-bonn.de

^cligeti@berkeley.edu

^dmpapucci@caltech.edu

^edrobinson@lbl.gov

by reversing detector effects, or they can be avoided by using dedicated MC samples for each theoretical model the measurement is confronted with. In this paper we present the newly developed tool, **Hammer** (*Helicity Amplitude Module for Matrix Element Reweighting*), designed expressly for the latter purpose.

Semitaquonic b hadron decays have long been known to be sensitive to new physics [1–7], and were first constrained at LEP [8]. At present, the measurements of the $R(D^{(*)})$ ratios show about a 3σ tension with SM predictions, when the D and D^* modes are combined [9]. In the future, much more precise measurements of semitaquonic decays are expected, not only for the $B \rightarrow D^{(*)}\tau\bar{\nu}$ channels, but also for the not yet studied decay modes, $\Lambda_b \rightarrow \Lambda_c\tau\bar{\nu}$, $B_s \rightarrow D_s^{(*)}\tau\bar{\nu}$, as well as involving excited charm hadrons in the final state.

All existing measurements of $R(D^{(*)})$ rely heavily on large MC simulations to optimize selections, provide fit templates in discriminating kinematic observables, and to model resolution effects and acceptances. Both the τ and the charm hadrons have short lifetimes and decay near the interaction point and measurements rely on reconstruction of the ensuing decay cascades. To reconstruct the decay products, often complex phase space cuts and detector efficiency dependencies come into play, and the measurement of the full decay kinematics is impossible due to the presence of multiple neutrinos. In addition, depending on the final state, a significant downfeed with similar experimental signatures from misreconstructed excited charm hadron states can be present. Isolation of semitaquonic decays from other background processes and the light-lepton final states, then requires precise predictions for the kinematics of the signal semitaquonic decay.¹ Often the limited size of the available simulated samples, required to account for all these effects, constitutes a dominant uncertainty of the measurements, see e.g. [10–12].

In the literature on the $R(D^{(*)})$ anomaly, it has become standard practice to reinterpret the experimental values of $R(D^{(*)})$ in terms of NP Wilson coefficients, even though all current ratio measurements were determined assuming the SM nature of semitaquonic decays. However, NP couplings generically alter decay distributions and acceptances. Therefore, they modify the signal and possibly background MC templates used in the extraction, and thus affect the measured values of $R(D^{(*)})$. This may introduce biases in NP interpretations: preferred regions and best-fit points for the Wil-

son coefficients can be incorrect; an instructive example of this is provided in Sec 2.3.

Consistent interpretations of the data with NP incorporated requires dedicated MC samples, ideally for each NP coupling value considered, which would permit directly fitting for the NP Wilson coefficients. This approach is sometimes referred to as ‘forward-folding’, and is naively a computationally prohibitively expensive endeavour. Such a program is further complicated because none of the MC generators current used by the experiments incorporate generic NP effects, nor do they include state-of-the-art treatments of hadronic matrix elements.

In this paper we present a new software tool, **Hammer**, that provides a solution to these problems: A fast and efficient means to reweight large MC samples to any desired NP, or to any description of the hadronic matrix elements. **Hammer** makes use of efficient amplitude-level and tensorial calculation strategies, and is designed to interface with existing experimental analysis frameworks, providing detailed control over which NP or hadronic descriptions should be considered. The desired reweighting can be implemented either in the event weights or in histograms of experimentally reconstructed quantities (both further discussed in Sec. 3). The only required input are the event-level truth-four-momenta of existing MC samples. Either the event weights and/or histogram predictions may be used, e.g., to generate likelihood functions for experimental fits. Some of the main ideas of **Hammer** were previously outlined in Refs. [13, 14].

In Sec. 2 we demonstrate the capabilities of **Hammer** by performing binned likelihood fits on mock measured and simulated data sets, that are created using the **Hammer** library, and corrected using an approximate detector response. In Sec. 3 a brief overview of the **Hammer** library and its capabilities are given. Section 4 provides a summary of our findings. Finally, Appendix A provides a detailed overview of the **Hammer** application programming interface.

2 New physics analyses

We consider two different analysis scenarios:

1. In order to explore what biases may arise in phenomenological studies if NP is present in Nature, we perform an illustrative $R(D^{(*)})$ toy measurement. This involves carrying out SM fits to mockups of measured data sets, that are generated for several different NP models. The recovered $R(D^{(*)})$ values are then compared to their actual NP values.

¹ Further complications arise from interference among the different spin states of the τ and among those of the charm hadron. Such effects have sometimes been neglected, treating the τ and charm hadron as stable particles, when simulations are corrected to account for more up-to-date hadronic models.

2. To demonstrate using a forward-folded analysis to assess NP effects without biases, we carry out fits to (combinations of) NP Wilson coefficients themselves, with either the SM or other NP present in the mock measured data sets.

The setting of these analyses is a B -factory-type environment. We focus on leptonic τ decays, but the procedures and results in this work are equally adaptable to the LHCb environment, and other τ decay modes or observables.

We emphasize that the derived sensitivities shown below are not intended to illustrate projections for actual experimental sensitivities *per se*. Such studies are better carried out by the experiments themselves.

2.1 MC sample

The input Monte Carlo sample used for our demonstration comprises four distinct sets of 10^5 events: one for each of the two signal cascades $B \rightarrow D(\tau \rightarrow e\nu\nu)\nu$, $B \rightarrow (D^* \rightarrow D\pi)(\tau \rightarrow e\nu\nu)\nu$ and for the two background processes, $B \rightarrow D e \nu$ and $B \rightarrow (D^* \rightarrow D\pi) e \nu$. These are generated with **EvtGen** R01-07-00 [15], using the Belle II beam energies of 7 GeV and 4 GeV. The second B meson decay, often used for identifying or ‘tagging’ the $B\bar{B}$ event and constraining its kinematic properties, are not included in the current analysis for simplicity, but can be incorporated in a **Hammer** analysis straightforwardly.

In each cascade, the $b \rightarrow c l \nu$ vertices are generated equidistributed in phase space (“pure phase space”) instead of using SM distributions. This reduces the statistical uncertainties that can otherwise arise from reweighting regions of phase space that are undersampled in the SM to NP scenarios in which they are not.²

2.2 Reweighting and fitting analysis

Hammer is used to reweight the MC samples into two-dimensional ‘NP generalized’ histograms (see Sec. 3), with respect to the reconstructed observables $|\mathbf{p}_\ell^*|$ and m_{miss}^2 , the light lepton momentum in the B rest frame and the total missing invariant mass of all neutrinos, respectively. Both variables are well-suited for separating signal from background decays involving only light leptons. In the cascade process of the leptonic τ decay in

$B \rightarrow D^{(*)}\tau\nu$, the signal lepton carries less momentum than the lepton from prompt $B \rightarrow D^{(*)}\ell\nu$ decays. Similarly, the missing invariant mass of $B \rightarrow D^{(*)}\ell\nu$ decays peaks strongly near $m_\nu^2 \simeq 0$, in contrast to $B \rightarrow D^{(*)}\tau\nu$ in which the multiple neutrinos in the final state permit large values of m_{miss}^2 .

The $B \rightarrow D^{(*)}$ processes are reweighted to the BLPR form factor parametrization [16], which includes predictions for NP hadronic matrix elements using HQET [17–20] at $\mathcal{O}(1/m_{c,b}, \alpha_s)$.

Charged particles are required to fall in the Belle II angular acceptance of 20° and 150° , and leptons are required to have a minimum kinetic energy of 300 MeV in the laboratory frame. An additional event weight is included to account for the slow pion reconstruction efficiencies from the $D^* \rightarrow D\pi$ decay, based on an approximate fit to the pion reconstruction efficiency curve from BaBar data [10, 21]. The analysis assumes that the second tagging B meson decay was reconstructed in hadronic modes, such that its four-momentum, $p_{B_{\text{tag}}}$, is accessible. In conjunction with the known beam four-momentum $p_{e^+e^-}$, the missing invariant mass can then be reconstructed as $m_{\text{miss}}^2 \equiv (p_{e^+e^-} - p_{B_{\text{tag}}} - p_{D^{(*)}} - p_\ell)^2$, and the four-momentum of the reconstructed lepton can be boosted into the signal B rest frame. A Gaussian smearing is added to the truth level m_{miss}^2 with a width of 0.5 GeV^2 to account for detector resolution and tagging- B reconstruction. No additional correction is applied to $|\mathbf{p}_\ell^*|$. Higher dimensional histograms including the reconstructed q^2 and the $D^* \rightarrow D\pi$ helicity angle may also be incorporated, but are omitted here for simplicity.

Hammer can be used to efficiently compute histograms for any given NP choice. The basis of NP operators is defined in Table 1, with respect to the Lagrangian

$$\mathcal{L} = \frac{4G_F}{\sqrt{2}} V_{cb} c_{XY} (\bar{c} \Gamma_X b) (\bar{\ell} \Gamma_Y \nu), \quad (2)$$

where $\Gamma_{X(Y)}$ is any Dirac matrix and c_{XY} is a Wilson coefficient. We shall generally write explicit Wilson coefficients as $c_{XY} = S_{qXY}, V_{qXY}, T_{qXY}$, where the S, V, T denotes the Lorentz structure, and $X, Y = L, R$ denotes the chirality. In this simplified analysis, we assume that NP only affects the $b \rightarrow c \tau \nu$ decays, and not the light-lepton modes.

In order to carry out Wilson coefficient fits, we wrap the **Hammer** application programming interface with a **gammaCombo** [22] compatible class. This allows one to use **Hammer**’s efficient reweighting of histogram bins to generate the relevant quantities required to calculate a likelihood function. We then carry out a fully two-dimensional binned likelihood fit in $|\mathbf{p}_\ell^*|$ and m_{miss}^2 , assuming Gaussian uncertainties. The fit uses 12×12 bins

² For an actual experimental analysis one would instead use **Hammer** to reweight SM MC samples. The correct statistical uncertainty of the reweighting can be incorporated, using weight squared uncertainties computed by the library. This information could be used, e.g., to adaptively generate additional pure phase space MC in undersampled regions.

Current	Label	Wilson Coefficient, c_{XY}	Operator
SM	SM	1	$[\bar{c}\gamma^\mu P_L b][\bar{\ell}\gamma_\mu P_L \nu]$
Vector	V_qLlL	V_{qLlL}	$[\bar{c}\gamma^\mu P_L b][\bar{\ell}\gamma_\mu P_L \nu]$
	V_qRlL	V_{qRlL}	$[\bar{c}\gamma^\mu P_R b][\bar{\ell}\gamma_\mu P_L \nu]$
	V_qLlR	V_{qLlR}	$[\bar{c}\gamma^\mu P_L b][\bar{\ell}\gamma_\mu P_R \nu]$
	V_qRlR	V_{qRlR}	$[\bar{c}\gamma^\mu P_R b][\bar{\ell}\gamma_\mu P_R \nu]$
Scalar	S_qLlL	S_{qLlL}	$[\bar{c}P_L b][\bar{\ell}P_L \nu]$
	S_qRlL	S_{qRlL}	$[\bar{c}P_R b][\bar{\ell}P_L \nu]$
	S_qLlR	S_{qLlR}	$[\bar{c}P_L b][\bar{\ell}P_R \nu]$
	S_qRlR	S_{qRlR}	$[\bar{c}P_R b][\bar{\ell}P_R \nu]$
Tensor	T_qLlL	T_{qLlL}	$[\bar{c}\sigma^{\mu\nu} P_L b][\bar{\ell}\sigma_{\mu\nu} P_L \nu]$
	T_qRlR	T_{qRlR}	$[\bar{c}\sigma^{\mu\nu} P_R b][\bar{\ell}\sigma_{\mu\nu} P_R \nu]$

Table 1 The $b \rightarrow c\ell\nu$ operator basis and coupling conventions. Also shown are the identifying Wilson coefficient labels used in **Hammer**. The normalization of the operators is as in Eq. (2).

$B \rightarrow D\tau\bar{\nu}$ Category	Fractions	Events / ab^{-1}
$B \rightarrow D\tau\bar{\nu}$	5.6%	800
$B \rightarrow D^*\tau\bar{\nu}$	2.3%	325
$B \rightarrow D\ell\bar{\nu}$	49.4%	7000
$B \rightarrow D^*\ell\bar{\nu}$	40.6%	5750
Irreducible background	2.0%	288
$B \rightarrow D^*\tau\bar{\nu}$ Category	Fractions	Events / ab^{-1}
$B \rightarrow D^*\tau\bar{\nu}$	5.4%	950
$B \rightarrow D^*\ell\bar{\nu}$	93.0%	16500
Irreducible background	1.6%	288

Table 2 The Asimov data set components. The fractions were motivated by Refs. [10, 21].

with equidistant bin widths for $|\mathbf{p}_\ell^*| \in (0.2, 2.2) \text{ GeV}$ and $m_{\text{miss}}^2 \in (-2, 10) \text{ GeV}^2$. The fits determine either $R(D^{(*)})$, or the real and imaginary parts of Wilson coefficients. The preferred SM coupling is determined simultaneously, in order to remove explicit dependence on $|V_{cb}|$.

We construct an Asimov data set [23] assuming the fractions and total number of events in Table 2, following from the number of events in Ref. [10, 21]. In the scans, the total number of events corresponds to an approximate integrated luminosity of 5 ab^{-1} of Belle II collisions. We assume events are reconstructed in two categories targeting $B \rightarrow D\tau\bar{\nu}$ and $B \rightarrow D^*\tau\bar{\nu}$. A fit for the real and imaginary parts of a single Wilson coefficient plus the (real) SM coupling thus has $2 \times 12 \times 12 - 3 = 285$ degrees of freedom.

A sizable downfeed background from D^* mesons misreconstructed as a D is expected in the $B \rightarrow D\tau\bar{\nu}$ channel via both the $B \rightarrow D^*\tau\bar{\nu}$ and $B \rightarrow D^*\ell\bar{\nu}$ de-

cays. This is taken into account by partitioning the simulated $B \rightarrow D^*\tau\bar{\nu}$ and $B \rightarrow D^*\ell\bar{\nu}$ events into two samples: One with the correct $m_{\text{miss}}^2 = (p_B - p_{D^*} - p_\ell)^2$ and the other with the misreconstructed $m_{\text{miss}}^2 = (p_B - p_D - p_\ell)^2$, which omits the slow pion. This downfeed reduces the sensitivity for the case that NP couplings induce opposite effects on the $B \rightarrow D\tau\bar{\nu}$ versus $B \rightarrow D^*\tau\bar{\nu}$ total rates or shapes. In addition to semileptonic processes, we assume the presence of an irreducible background from secondaries (i.e., leptons from semileptonic D meson decays), fake leptons (i.e., hadrons that were misidentified as leptons) and semileptonic decays from higher charm resonances (i.e., D^{**} states). The irreducible background is modeled in a simplified manner by assuming 10 background events in each of the 12×12 bins, totaling overall 1440 events per category.

Figure 1 shows the impact on the fit variables of three benchmark models that we use to investigate the effects of new physics:

- i) The R_2 leptoquark model, which sets $S_{qLlL} \simeq 8 T_{qLlL}$ (see, e.g., Refs. [24, 25]);
- ii) A pure tensor model, via T_{qLlL} ;
- iii) A right-handed vector model, via V_{qRlL} .

For the ratio plots in Fig. 1, we fix the NP Wilson coefficients to specific values to illustrate the shape changes they induce in $|\mathbf{p}_\ell^*|$ and m_{miss}^2 . The R_2 leptoquark model and tensor model exhibit sizable shape changes. The right-handed vector model shows only an overall normalization change for $B \rightarrow D\tau\bar{\nu}$, with no change in shape compared to the SM, because the axial-vector $B \rightarrow D$ hadronic matrix element vanishes by parity and angular momentum conservation. For $B \rightarrow D^*$, both vector and axial vector matrix elements are nonzero, so that introducing a right-handed vector current leads to shape and normalization changes.

Figure 2 shows the projections of the constructed Asimov data set, as well as the distributions expected for the three NP models. The latter have the same couplings as those shown in Fig. 1.

2.3 $R(D^{(*)})$ biases from new physics truth

Many NP analyses and global fits to the $R(D^{(*)})$ measurements – together with other potentially template-sensitive observables, including q^2 spectra – have been carried out by a range of phenomenological studies (see, e.g., Refs. [24–32]). As mentioned above, the standard practice has been to fit NP predictions to the world-average values of $R(D^{(*)})$ (and other data) to determine confidence levels for allowed and excluded NP couplings. However, because the $R(D^{(*)})$ measurements use SM-based templates, and because the presence of

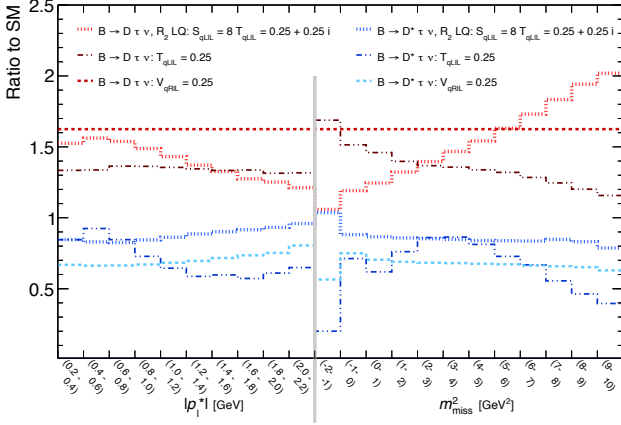


Fig. 1 The ratios of various Wilson coefficient working points to the SM in $|\mathbf{p}_\ell^*|$ and m_{miss}^2 are shown. The coupling strengths assumed are shown in the legend. For more details see text.

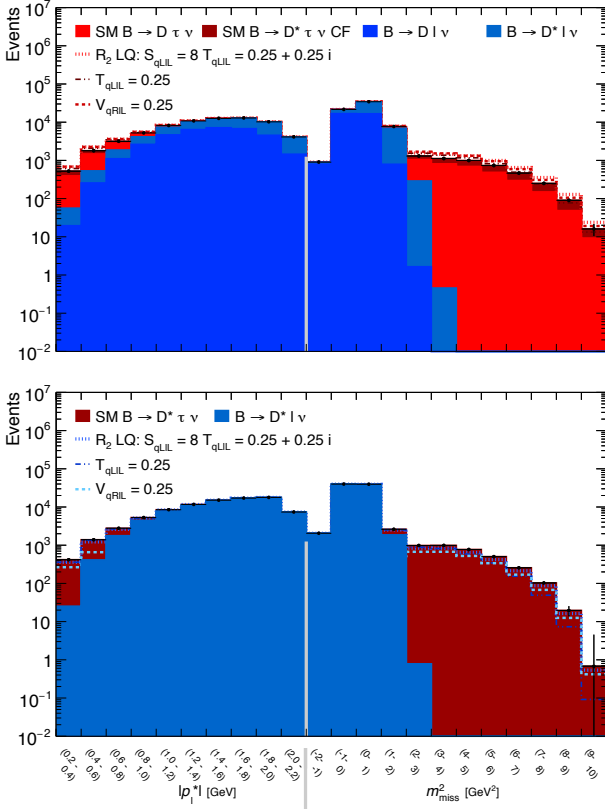


Fig. 2 The $B \rightarrow D\tau\bar{\nu}$ (top) and $B \rightarrow D^*\tau\bar{\nu}$ (bottom) distributions in $|\mathbf{p}_\ell^*|$ and m_{miss}^2 in the Asimov data set. The number of events correspond to an estimated number of reconstructed events at Belle II with 5 ab^{-1} .

NP operators can strongly alter acceptances and kinematic distributions, such analyses can lead to incorrect best-fit values or exclusions of NP Wilson coefficients.

To illustrate such a bias, we fit SM MC templates to NP Asimov data sets, that are generated with **Hammer** for two different NP ‘truth’ benchmark points: the 2HDM Type II with $S_{qRIL} = -2$, corresponding to

$\tan\beta/m_{H^\pm} \simeq 0.5\text{ GeV}^{-1}$; and the R_2 leptoquark model with $S_{qLIL} = 8T_{qLIL} = 0.25 + 0.25i$. (These models and couplings are for illustration and are arbitrary choices; our goal here is only to demonstrate the generic biases that can occur.) We replicate the fit of all existing measurements, allowing the normalizations of the D and D^* modes (and the light leptonic final states) to float independently, without imposing e.g. their predicted SM relationship. This fit leads to a best-fit ellipse in the $R(D) - R(D^*)$ plane.

In Fig. 3 we show the recovered values, $R(D^{(*)})_{\text{rec}}$, obtained from this procedure, and compare them to the actual predictions of the given NP truth benchmark point, $R(D^{(*)})_{\text{th}}$. For ease of comparison, we normalize the $R(D^{(*)})$ values against the SM predictions for $R(D^{(*)})_{\text{SM}}$. For both NP models, the recovered ratios from fitting the Asimov data set exclude the truth $R(D^{(*)})_{\text{th}}$ values at $\gtrsim 4\sigma$. This illustrates the sizeable bias in the measured $R(D^{(*)})$ values that ensues from carrying out fits with an SM template, if instead NP actually contributes to the measurements.

We also show in Fig. 3 the equivalent bias arising from a naïve fit of the $R(D^{(*)})$ NP prediction that attempts to recover the complex Wilson coefficient. This is done by parametrizing $R(D^{(*)})_{\text{th}} = R(D^{(*)})[c_{XY}]$, and fitting this expression to the recovered $R(D^{(*)})_{\text{rec}}$ values. Explicitly, one calculates CLs in the Wilson coefficient space via the 2 degree of freedom chi-square $\chi^2 = \mathbf{v}^T \sigma_{R(D^{(*)})}^{-1} \mathbf{v}$, with $\mathbf{v} = (R(D)_{\text{th}} - R(D)_{\text{rec}}, R(D^*)_{\text{th}} - R(D^*)_{\text{rec}})$. The resulting best fit Wilson coefficient regions similarly exclude the truth values.

Thus, the allowed or excluded regions of NP couplings determined from fits to the $R(D^{(*)})$ measurements must be treated with caution, as these fits do not include effects of the NP distributions in the MC templates. Similarly, results of global fits should be interpreted carefully when assessing the level of compatibility with specific NP scenarios.

2.4 New physics Wilson coefficient fits

Instead of considering observables like $R(D^{(*)})$, for phenomenological studies to be able to properly make interpretations and test NP models, experiments should provide direct constraints on NP Wilson coefficients themselves. For example, this could be done with simplified likelihood ratios that profile out all irrelevant nuisance parameters from, e.g., systematic uncertainties or information from sidebands or control channels, or by other means.

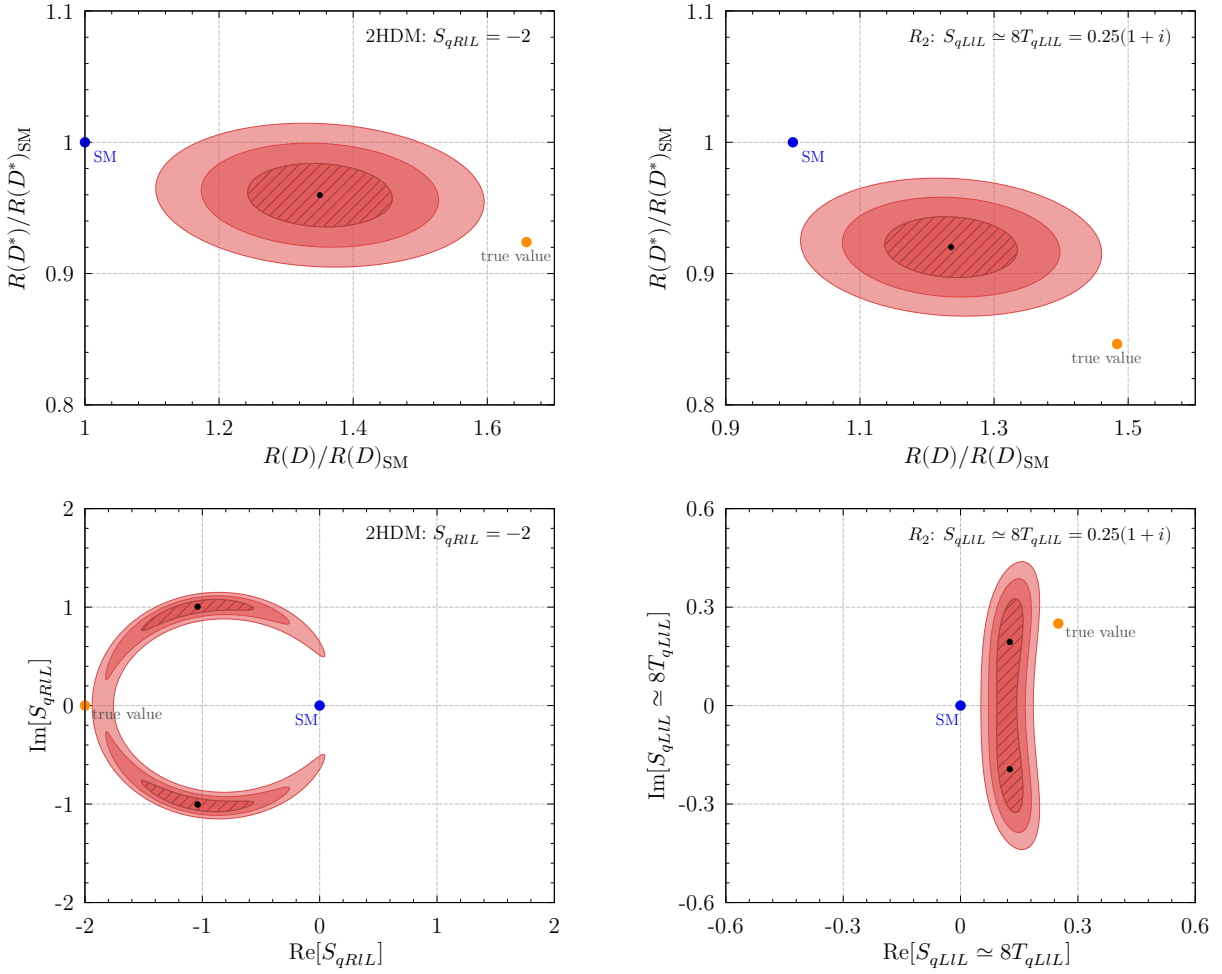


Fig. 3 *Top:* Illustrations of biases from fitting an SM template to two NP ‘truth’ benchmark models: the 2HDM type II with $S_{qRLL} = -2$ (left) and the R_2 leptiquark model with $S_{qLIL} = 8T_{qLIL} = 0.25 + 0.25i$ (right). The black dot corresponds to the predicted ‘truth value’ of $R(D^*)$ for the NP model, to be compared to the 68%, 95% and 99% CLs of the SM fit to the NP Asimov data sets, shown in shades of red. The resulting recovered best fit ratios for the 2HDM type II fit are $R(D)/R(D)_{\text{SM}} = 1.35 \pm 0.07$ (predicted: 1.66) and $R(D^*)/R(D^*)_{\text{SM}} = 0.96 \pm 0.02$ (predicted: 0.92). The R_2 leptiquark model best fit ratios are $R(D)/R(D)_{\text{SM}} = 1.24 \pm 0.07$ (predicted: 1.48) and $R(D^*)/R(D^*)_{\text{SM}} = 0.92 \pm 0.02$ (predicted: 0.85). *Bottom:* The best fit regions for the 2HDM and R_2 model Wilson coefficients obtained from fitting $R(D^*)$ NP predictions to the recovered $R(D^*)$ CLs for each NP model. The shades of red denote CLs as in the top row, and the best fit [truth] Wilson coefficients are shown by black [orange] dots.

As an example, we now use **Hammer** to perform such a fit for the real and imaginary parts of the NP Wilson coefficients, using the set of three NP models in Sec. 2.2 as templates. These are fit to the same two truth benchmark scenarios as in Fig. 4: a truth SM Asimov data set; and a truth Asimov data set reweighted to the 2HDM Type II with $S_{qRLL} = -2$.

Figure 4 shows in shades of red the 68%, 95% and 99% confidence levels (CLs) of the three NP model scans of SM Asimov data sets. For the SM truth benchmark, the corresponding best fit points are always at zero NP couplings. The derived CLs then correspond to the expected median exclusion of the fitted NP coupling under the assumption the SM is true.

We further show in shades of yellow the same fit CLs for the 2HDM truth benchmark Asimov data set. These latter fits illustrate a scenario in which NP is present, but is analyzed with an incomplete or incorrect set of NP Wilson coefficients. Depending on the set of coefficients, we see from the $\Delta\chi^2$ of the best fit points that the new physics might be obfuscated or wrongly identified. This underlines the importance for LHCb and Belle II to eventually carry out an analysis with the full set of operators listed in Table 1.

3 The Hammer library

In this section we present core interface features and calculational strategies of the **Hammer** library. Details

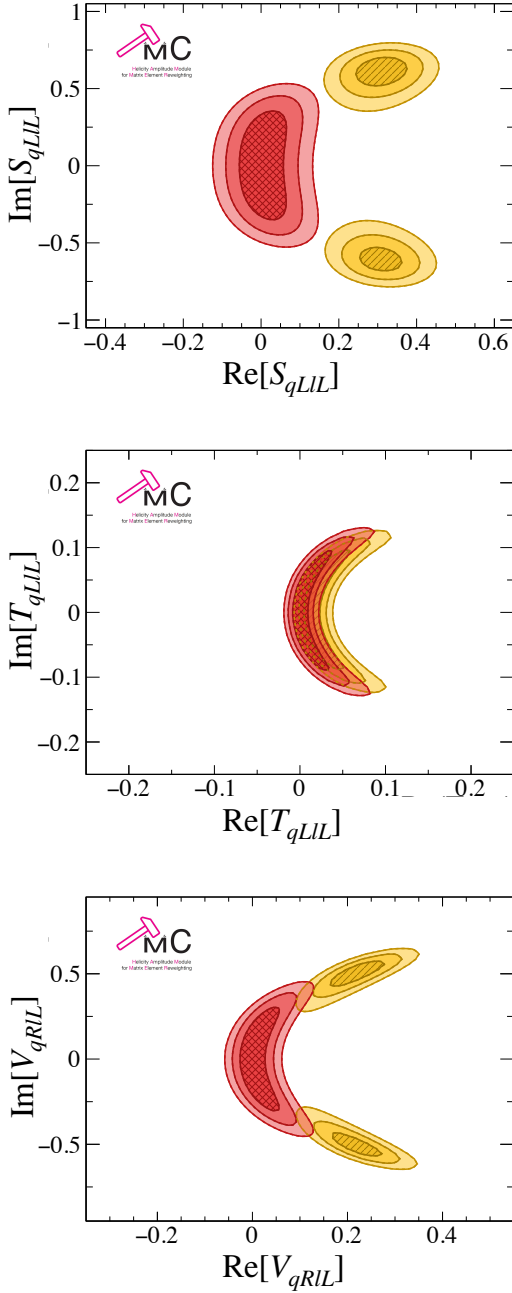


Fig. 4 The 68%, 95% and 99% CL allowed regions of three models under consideration, from fitting the SM (red) and 2HDM type II (yellow and with $S_{qRL} = -2$) Asimov data sets. (Top) R_2 leptokark model with $S_{qLIL} = 8T_{qLIL}$; (middle) NP in the form of a left-handed tensor coupling; (bottom) NP in the form of a right-handed vector coupling.

of the code structure, implementation, and use, can be found in the **Hammer** manual [33]; here we provide only an overview.

3.1 Reweighting

We consider an MC event sample, comprising a set of events indexed by I , with weights w_I and truth-level kinematics $\{q\}_I$. Reweighting this sample from an ‘old’ to a ‘new’ theory requires the truth-level computation of the ratio of the differential rates

$$r_I = \frac{d\Gamma_I^{\text{new}}/d\mathcal{PS}}{d\Gamma_I^{\text{old}}/d\mathcal{PS}}, \quad (3)$$

applied event-by-event via the mapping $w_I \mapsto r_I w_I$. The ‘old’ or ‘input’ or ‘denominator’ theory is typically the SM plus (where relevant) a hadronic model — that is, a form factor (FF) parametrization. (It may also be composed of pure phase space (PS) elements, see App. A.2.) The ‘new’ or ‘output’ or ‘numerator’ theory may involve NP beyond the Standard Model, or a different hadronic model, or both.

Historically, the primary focus of the library is reweighting of $b \rightarrow c\ell\nu$ semileptonic processes, often in multistep cascades such as $B \rightarrow D^{(*)}(\rightarrow DY)\tau(\rightarrow X\nu)\bar{\nu}$. However, the library’s computational structure is designed to be generalized beyond these processes, and we therefore frame the following discussion in general terms, before returning to the specific case of semileptonic decays.

3.2 New Physics generalizations

The **Hammer** library is designed for the reweighting of processes via theories of the form

$$\mathcal{L} = \sum_{\alpha} c_{\alpha} \mathcal{O}_{\alpha}. \quad (4)$$

where \mathcal{O}_{α} are a basis of operators, and c_{α} , are SM or NP Wilson coefficients (defined at a fixed physical scale; mixing of the Wilson coefficients under RG evolution, if relevant, must be accounted for externally to the library). We specify in Table 1 the conventions used for various $b \rightarrow c\ell\nu$ four-Fermi operators and other processes included in the library.

The corresponding process amplitudes may be expressed as linear combinations $c_{\alpha}\mathcal{A}_{\alpha}$. They may also be further expressed as a linear sum with respect to a basis of form factors, F_i , that encode the physics of hadronic transitions (if any).³ In general, then, an amplitude may

³In all $b \rightarrow c$ processes currently handled by **Hammer** (see Table 3 for a list) the form factors are functions of $q^2 = (p_{H_b} - p_{H_c})^2$, or equivalently of the dimensionless kinematic variable,

$$w = v \cdot v' = \frac{m_{H_b}^2 + m_{H_c}^2 - q^2}{2m_{H_b}m_{H_c}}, \quad (5)$$

be written in the form

$$\mathcal{M}^{\{s\}}(\{q\}) = \sum_{\alpha,i} c_{\alpha} F_i(\{q\}) \mathcal{A}_{\alpha i}^{\{s\}}(\{q\}), \quad (6)$$

in which $\{s\}$ are a set of external quantum numbers and $\{q\}$ the set of four-momenta.⁴ The object $\mathcal{A}_{\alpha i}$ is an NP- and FF-generalized *amplitude tensor*. In the case of cascades, relevant for $B \rightarrow D^{(*,**)}(\rightarrow DY) \tau(\rightarrow X\nu)\bar{\nu}$ decays, the amplitude tensor may itself be the product of several subamplitudes, summed over several sets of internal quantum numbers. The corresponding polarized differential rate

$$\begin{aligned} \frac{d\Gamma^{\{s\}}}{d\mathcal{PS}} &= \sum_{\alpha,i,\beta,j} c_{\alpha} c_{\beta}^{\dagger} F_i F_j^{\dagger}(\{q\}) \mathcal{A}_{\alpha i}^{\{s\}} \mathcal{A}_{\beta j}^{\dagger\{s\}}(\{q\}), \\ &= \sum_{\alpha,i,\beta,j} c_{\alpha} c_{\beta}^{\dagger} F_i F_j^{\dagger}(\{q\}) \mathcal{W}_{\alpha i \beta j}, \end{aligned} \quad (7)$$

in which the phase space differential form $d\mathcal{PS}$ includes on-shell δ -functions and geometric or combinatoric factors, as appropriate.

The outer product of the amplitude tensor, defined as $\mathcal{W} \equiv \mathcal{A}\mathcal{A}^{\dagger}$, is a *weight tensor*. The object $\sum_{i,j} F_i F_j^{\dagger} \mathcal{W}_{\alpha i \beta j}$ in Eq. (7) is independent of the Wilson coefficients: Once this object is computed for a specific $\{q\}$ – an event – it can be contracted with any choice of NP to generate an event weight. Similarly, on a patch of phase space Ω – e.g., the acceptance of a detector or a bin of a histogram – the marginal rate can now be written as

$$\Gamma_{\Omega}^{\{s\}} = \sum_{\alpha,\beta} c_{\alpha} c_{\beta}^{\dagger} \int_{\Omega} d\mathcal{PS} \sum_{ij} F_i F_j^{\dagger}(\{q\}) \mathcal{W}_{\alpha i \beta j}^{\{s\}}(\{q\}). \quad (8)$$

The Wilson coefficients factor out of the phase space integral, so that the integral itself generates a NP-generalized tensor. After it is computed once, it can be contracted with any choice of NP Wilson coefficients, c_{α} , thereafter.

The core of **Hammer**’s computational philosophy is based on the observation that this contraction is computationally much more efficient than the initial computation (and integration). Hence efficient reweighting is achieved by

- Computing NP (and/or FF, see below) generalized objects, and storing them;

with four velocities $v = p_{H_b}/m_{H_b}$ and $v' = p_{H_c}/m_{H_c}$. For decays with multi-hadron final states, such as the $\tau \rightarrow n\pi$, $n \geq 3$, the form factors are also dependent on multiple invariant masses of the final state hadrons. Thus, $b \rightarrow c\tau\nu$ decays followed by with subsequent hadronic τ decays involve at least two separate sets of hadronic functions at the amplitude level.

⁴ The momenta of an event passed to the library must all be specified in the same frame. The choice of frame is arbitrary.

Process	Form factor parametrizations
$B \rightarrow D^{(*)}\ell\nu$	ISGW2* [34, 35], BGL* [36–38], CLN* [‡] [39], BLPR [‡] [16]
$B \rightarrow (D^* \rightarrow D\pi)\ell\nu$	ISGW2*, BGL* [‡] , CLN* [‡] , BLPR [‡]
$B \rightarrow (D^* \rightarrow D\gamma)\ell\nu$	ISGW2*, BGL* [‡] , CLN* [‡] , BLPR [‡]
$\tau \rightarrow \pi\nu$	—
$\tau \rightarrow \ell\nu\nu$	—
$\tau \rightarrow 3\pi\nu$	RCT* [40–42]
$B \rightarrow D_0^*\ell\nu$	ISGW2*, LLSW* [43, 44], BLR [‡] [45, 46]
$B \rightarrow D_1^*\ell\nu$	ISGW2*, LLSW*, BLR [‡]
$B \rightarrow D_1\ell\nu$	ISGW2*, LLSW*, BLR [‡]
$B \rightarrow D_2^*\ell\nu$	ISGW2*, LLSW*, BLR [‡]
$\Lambda_b \rightarrow \Lambda_c \ell\nu$	PCR* [47], BLRS [‡] [48, 49]
Planned for next release	
$B_{(c)} \rightarrow \ell\nu$	MSbar
$B \rightarrow (\rho \rightarrow \pi\pi)\ell\nu$	BCL*, BSZ
$B \rightarrow (\omega \rightarrow \pi\pi\pi)\ell\nu$	BCL*, BSZ
$B_c \rightarrow (J/\psi \rightarrow \ell\ell)\ell\nu$	
$\Lambda_b \rightarrow \Lambda_c^* \ell\nu$	PCR*, BLRS
$\tau \rightarrow 4\pi\nu$	RCT*
$\tau \rightarrow (\rho \rightarrow \pi\pi)\nu$	—

Table 3 Presently implemented amplitudes in the **Hammer** library, and corresponding form factor parametrizations. SM-only parametrizations are indicated by a * superscript. Form factor parametrizations that include linearized variations are denoted with a ‡ superscript. These are named in the library by adding a “Var” suffix, e.g. “BGLVar”.

- Contracting them thereafter for any given NP (and/or FF) choice to quickly generate a desired NP (and/or FF) weight.

3.3 Form factor generalizations

Similarly to the NP Wilson coefficients, it is often desirable to be able to vary the FF parameterizations themselves. For instance, one might contemplate variations along the error eigenbasis of a fit to the FF parameters, or FF parametrizations that are linearized with respect to a basis of parameters, such as the BGL parametrization [36–38] in $B \rightarrow D^{(*)}\ell\nu$. To this end, an FF parametrization with a parameter set $\{\mu\}$ can be linearized around a (best-fit) point, $\{\mu^0\}$ so that

$$F_i(\{q\}; \{\mu\}) = F_i(\{q\}, \{\mu^0\}) + \sum_a F'_{i,a}(\{q\}, \{\mu^0\}) e_a, \quad (9)$$

where ‘a’ is one or more *variational indices* and e_a is the variation. In the language of the error eigenbasis case, $F'_{i,a}$ is the perturbation of F_i in the a th principal component e_a of the parametric fit correlation matrix.

Defining $\xi_a \equiv (1, e_a)$ and $\Phi_{i,a+1} \equiv (F_i, F'_{i,a})$, so that Eq. (9) becomes

$$\sum_a \xi_a \Phi_{i,a} = F_i + \sum_{a'} F'_{i,a'} e_{a'}, \quad (10)$$

then the differential rate

$$\begin{aligned} \frac{d\Gamma^{\{s\}}}{d\mathcal{PS}} &= \sum_{\alpha, a, \beta, b} c_\alpha c_\beta^\dagger \xi_a \xi_b^\dagger \mathcal{U}_{\alpha\beta b}^{\{s\}}, \\ \mathcal{U}_{\alpha\beta b}^{\{s\}} &\equiv \sum_{ij} \Phi_{i,a} \Phi_{j,b}^\dagger(\{q\}) \mathcal{W}_{\alpha\beta j}^{\{s\}}(\{q\}), \end{aligned} \quad (11)$$

with \mathcal{U} an NP- and FF-generalized weight tensor. The ξ_a are independent of $\{q\}$ and factor out of any phase space integral just as the Wilson coefficients do. That is, an integral on any phase space patch,

$$\Gamma_\Omega^{\{s\}} = \sum_{\alpha, \beta, a, b} c_\alpha c_\beta^\dagger \xi_a \xi_b^\dagger \int_\Omega d\mathcal{PS} \mathcal{U}_{\alpha\beta b}^{\{s\}}. \quad (12)$$

One may thus tensorialize the amplitude with respect to Wilson coefficients and/or FF linearized variations, to be contracted later with NP or FF variation choices (the latter within the regime of validity of the FF linearization). Hereafter, the ξ_a are referred to as ‘FF uncertainties’ or ‘FF eigenvectors’ following the nominal fit correlation matrix example.

3.4 Rates

In certain cases, it is also useful to compute and fold in an overall ratio of rates $\Gamma^{\text{old}}/\Gamma^{\text{new}}$, or the rates themselves, $\Gamma^{\text{new,old}}$, may be required. For example, if the MC sample has been initially generated with a fixed overall branching ratio, \mathcal{B}_{new} , one might wish to enforce this constraint via an additional multiplicative factor $\mathcal{B}_{\text{old}}/\mathcal{B}_{\text{new}}$.

The different components computed by **Hammer** are then:

- (i) The NP- and/or FF-generalized tensor for $(d\Gamma_I^{\text{new}}/d\mathcal{PS})/(d\Gamma_I^{\text{old}}/d\mathcal{PS})$, via Eq. (11), noting that the denominator carries no free NP or FF variational index. (The ratio r_I is then itself generally at least a rank-2 tensor.);
- (ii) The NP- and/or FF-generalized *rate tensors* $\Gamma^{\text{old, new}}$, which need be computed only once for an entire sample. (These rates require integration over the phase space, which is achieved by a dedicated multidimensional Gaussian quadrature integrator.)

3.5 Primary code functionalities

The calculational core of **Hammer** computes the NP or FF generalized tensors event-by-event for any process known to the library (see Table 3 for a list), and as specified by initialization choices (more detail is provided in Sec. 3.6) and specified form factor parametrizations. This core is supplemented by a wide array of functionalities to permit manipulation the resulting NP- and FF-generalized weight tensors as needed. This may include binning — equivalent to integrating on a phase space patch — the weight tensors into a histogram of any desired reconstructed observables, and/or it may include folding of detector simulation smearings, etc. Such histograms have NP- and FF-generalized tensors as bin entries, and we therefore call them *generalized* or *tensor* histograms. Once such NP- and FF-generalized tensor objects are computed and stored, contraction with NP or FF eigenvector choices permits the library to efficiently generate actual event weights or histogram bin weights for any theory of interest.

The architecture of **Hammer** is designed around several primary functionalities:

1. Provide an interface to determine which processes are to be reweighed, and which (possibly multiple) schemes for form factor parametrizations are to be used. This includes handling for (sub)processes that were generated as pure phase space.
2. Parse events into cascades of amplitudes known to the library, and compute their corresponding NP- and/or FF-generalized amplitude or weight tensor, as well as the respective rate tensors, as needed.
3. Provide an interface to generate histograms (of arbitrary dimension), and bin the event weight tensors — i.e., $r_I w_I$, as in Eq. (3) — into these histograms, as instructed. This includes functionality for weight-squared statistical errors, functionality for generation of ROOT histograms, as well as extensive internal architecture for efficient memory usage.
4. Efficiently contract generalized weight tensors or bin entries against specific FF variation or NP choices, to generate an event or bin weight. This includes extensive internal architecture to balance speed versus memory requirements.
5. Provide interface to save and reload amplitude or weight tensors or generalized histograms, to permit quick reprocessing into weights from precomputed or ‘initialized’ tensor objects.

Examples of the implementation of these functionalities are shown in many examples provided with the source code.

3.6 Code flow

A **Hammer** program may have two different types of structure: An *initialization* program, so called as it runs on MC as input, and may generate **Hammer** format files; or a *analysis* program, which may reprocess histograms or event weights that have already been saved in an initialization run. Pertinent details of the elements of the application programming interface mentioned below are provided in Appendix A, with more details in the **Hammer** manual.

An initialization program has the generic flow:

1. Create a **Hammer** object.
2. Declare included or forbidden processes, via `includeDecay` and `forbidDecay`.
3. Declare form factor schemes, via `addFFScheme` and `setFFInputScheme`.
4. (Optional) Add histograms, via `addHistogram`.
5. (Optional) Declare the MC units, via `setUnits`.
6. Initialize the **Hammer** class members with `initRun`.
7. (Optional) Change FF default settings with `setOptions`, or (if not SM) declare the Wilson coefficients for the input MC via `setWilsonCoefficients`.
8. (Optional) Fix Wilson coefficient (Wilson coefficient and/or FF uncertainty) choice to special choices in weight calculations (histogram binnings), via `specializeWCInWeights` (`specializeWCInHistogram` and/or `specializeFInHistogram`).
9. Each event may contain multiple processes, e.g., a signal and tag B decay. Looping over the events:
 - (a) Initialize event with `initEvent`. For each process in the event:
 - i. Create a **Hammer** `Process` object.
 - ii. Add particles and decay vertices to create a process tree, via `addParticle` and `addVertex`.
 - iii. Decide whether to include or exclude processes from an event via `addProcess` and/or `removeProcess`.
 - (b) Compute or obtain event observables – specific particles can be extracted with `getParticlesByVertex` or other programmatic means – and specify the corresponding histogram bins to be filled via `fillEventHistogram`.
 - (c) Initialize and compute the process amplitudes and weight tensors for included processes in the event, and fill histograms with event tensor weights – the direct product of include process tensor weights – via `processEvent`.
 - (d) (Optional) Save the weight tensors for each event, with `saveEventWeights` to a buffer.
10. (Optional) Generate histograms with `getHistogram(s)` and/or save them with `saveHistograms`. NP choices are implemented

with `setWilsonCoefficients`, FF variations are set with `setFFEigenvalues`.

11. (Optional) Save the rate tensors, with `saveRates` to a buffer.
12. (Optional) Save an autogenerated **bibTeX** list of references used in the run with `saveReferences`.

By contrast, an *analysis* program (from a previously initialized sample, stored in a buffer) has the generic flow:

1. Create a **Hammer** object and specify the input file.
2. Load or merge the run header — include or forbid specifications, FF schemes, or histograms — with `loadRunHeader` (after `initRun`). One may further declare additional histograms to be compiled (from saved event weight data) via `addHistogram`.
3. (Optional) Load or merge saved histograms with `loadHistograms`, and/or generate desired histograms with `getHistogram(s)`. NP choices are implemented with `setWilsonCoefficients`.
4. (Optional) Looping over the events:
 - (a) Initialize event with `initEvent`.
 - (b) If desired, remove processes from an event with `removeProcess`.
 - (c) Reload event weights with `loadEventWeights`.
 - (d) Specify histograms to be filled via `fillEventHistogram`.
 - (e) Fill histograms with event weights via `processEvent`.

4 Conclusions

Precision measurements of $b \rightarrow c\tau\bar{\nu}$ decays require large Monte Carlo samples, which incorporate detailed simulations of detector responses and physics backgrounds. The limited statistics due to the computational cost of these simulations are often a leading systematic uncertainty in the measurements, and it is prohibitively expensive to generate fully simulated MC samples for arbitrary NP models or descriptions of hadronic matrix elements.

In this paper we described the **Hammer** library, and illustrated its utility. **Hammer** allows the fast and efficient reweighting of existing SM (or phase-space based) MC samples to arbitrary NP models. In addition, **Hammer** can be used to change form factor parametrizations and/or incorporate uncertainties from form factors into experimental measurements. **Hammer** provides a computationally fast way for binned fits to generate predictions, and we implement a demonstrative forward-folding fit to constrain NP Wilson coefficients using this feature. Such a fit should be carried out by experimental

collaborations in future measurements to provide reliable constraints on NP contributions in semileptonic $b \rightarrow c\tau\bar{\nu}$ decays. The results will allow people outside the collaborations to make correct interpretations of the data, which has not been possible to date without potentially sizeable biases. To demonstrate this latter point, we carried out toy NP analyses using SM fits to NP Asimov data sets, and showed that sizeable biases can indeed occur. **Hammer** is open source software and we are looking forward to the experimental results and interpretations it will enable.

Acknowledgements **Hammer** has been developed with the active participation and testing by many colleagues. We especially thank from LHCb Julián García Pardiñas, Lucia Grillo, Donal Hill, Simone Meloni, Adam Morris, Patrick Owen, and Luke Scantlebury-Smead, for their extensive feedback, discussions, questions, and beta testing during development. We similarly thank from Belle II Kilian Lieret, Thomas Lueck, Felix Metzner, Markus Prim, and Maximilian Welsch. We thank David Shih for discussions and comments on the manuscript. Thanks are also due to all interested users on Belle, Belle II, BaBar, LHCb, and CMS, for many helpful discussions, questions, testing, and feedback. FB was supported by the DFG Emmy-Noether Grant No. BE 6075/1-1. SD was supported by the German Ministry of Research and Science (BMBF). ZL, MP and DR were supported in part by the Office of High Energy Physics of the U.S. Department of Energy under contract DE-AC02-05CH11231. We thank the Aspen Center of Physics, supported by the NSF grant PHY-1607611, where parts of this work were completed. This work also used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. FB thanks Kim Scott, Bob Michaud and Julie Michaud-B for their hospitality, many good discussions and in general a great time in Houston, where part of this paper was written.

A: Core elements of the Application Programming Interface

The user interface of the **Hammer** library provides four main classes: the **Hammer** class itself; the **Process** and **Particle** classes, used to create events; and the **IOBuffer** class used for saving and loading precomputed objects. A schematic of the architecture of **Hammer** is shown in Fig. 5.

In the following we describe various core parts of the **Hammer** Application Programming Interface (API), with many more details available in the code manual. The library itself is implemented in C++, along with a Python3 wrapper of the API; we will consider here the C++ interface only. This discussion is ordered by scope, rather than the typical code flow. Further details can be found in the **Hammer** manual [33].

A.1: Building processes and events

A typical decay cascade is contained in the library by the **Process** class; an event may contain multiple **Process** instances as e.g., is the case for a signal plus tag $B\text{--}\bar{B}$ pair. Each cascade may be simply represented in graphical terms as a ‘process tree’, as shown in Fig. 6: Each particle in the cascade is assigned an index, and each decay vertex is represented as a map from a parent index, to the indices of all its daughters. **Hammer** assembles the process tree through two methods **Process::addParticle** and **Process::addVertex**. The former adds a **Particle** class object – a momentum and a PDG code – to a container of particles; the latter fills the map of each parent index to its daughters for each decay vertex.

In the case of Fig. 6, the first two vertices of the cascade may be built explicitly as follows:

```
Process proc;
size_t idx0 = proc.addParticle(Particle{{E_0,
    px_0, py_0, pz_0}, pdg_0});
size_t idx1 = proc.addParticle(Particle{{E_1,
    px_1, py_1, pz_1}, pdg_1});
size_t idx2 = proc.addParticle(Particle{{E_2,
    px_2, py_2, pz_2}, pdg_2});
size_t idx3 = proc.addParticle(Particle{{E_3,
    px_3, py_3, pz_3}, pdg_3});
size_t idx7 = proc.addParticle(Particle{{E_7,
    px_7, py_7, pz_7}, pdg_7});
size_t idx8 = proc.addParticle(Particle{{E_8,
    px_8, py_8, pz_8}, pdg_8});

proc.addVertex(idx0, {idx1,idx2,idx3});
proc.addVertex(idx2, {idx7,idx8});
```

and so on. Particles and vertices need not be added in order; helper functions are provided in the code examples for automatically parsing HepMC files.

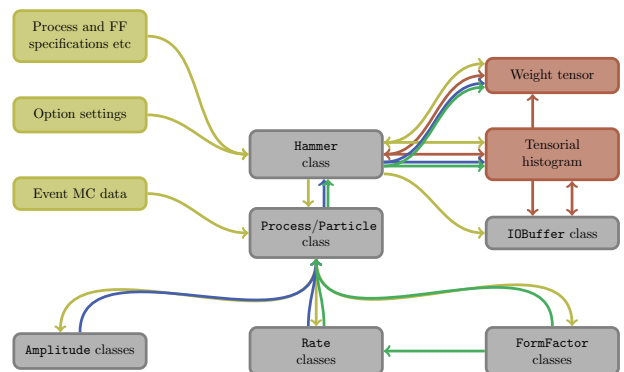


Fig. 5 Schematic architecture of **Hammer**. The flow of user specified choices or event data is shown by yellow arrows. Blue (green) arrows denote the flow of calculational information, in particular amplitude, weight or rate (form factor) tensors. Red arrows highlight the flow of **Hammer** output, which may be saved or reloaded. Most internal **Hammer** classes are not shown in this schematic.

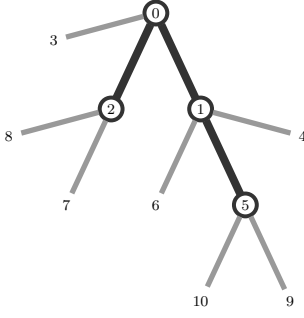


Fig. 6 Example process tree for a decay cascade involving 10 particles (numbers), 4 vertices (circles) and 3 edges (dark lines).

A.2: Specifications

The **Hammer** library contains an interpreter that maps a string representation of a vertex – a vertex string – to all possible charge conserving processes allowed by the specified particle names. The interpreter uses the syntax that particle names are parsed by a capital letter: the full list of names is provided in the manual. (For example the vertex string `"D*DPi"` is interpreted as all twelve possible $D^* \rightarrow D\pi$ vertices, while `"D*+DPi"` is interpreted as only the $D^{*+} \rightarrow D^+\pi^0$, $D^{*+} \rightarrow D^0\pi^+$, and (the heavily CKM suppressed) $D^{*+} \rightarrow \bar{D}^0\pi^+$ decay.)

The decay processes to be reweighed by **Hammer** are specified via `Hammer::includeDecay`, which takes a single vertex string or vector of vertex strings $\{V_1, V_2, \dots, V_n\}$ as an argument, and may be invoked multiple times. Each `includeDecay` specification is *inclusive* and permits any process tree whose full set of vertices contains all of $\{V_1, V_2, \dots, V_n\}$. The boolean logic applied by `includeDecay` is AND between each vertex string element, and OR between separate invocations of `includeDecay`. For example

```
Hammer ham;
ham.includeDecay({"BD*TauNu", "D*DGamma"});
ham.includeDecay({"BDMuNu"});
```

means ‘Reweight a process that either contains vertices ($B \rightarrow D^*\tau\nu$ and $D^* \rightarrow D\gamma$) or the vertex ($B \rightarrow D\mu\nu$)’. Hence, e.g., $\bar{B}^0 \rightarrow (D^{*+} \rightarrow (D^+ \rightarrow K^+\pi^+\pi^-\gamma)(\tau^- \rightarrow \ell^-\nu\nu))$ would be included. Recombination of radiative photons (produced during MC generation by **PHOTOS**) is handled automatically by the library, and need not be specified in `includeDecay` specifications. Processes may instead be forbidden with the `Hammer::forbidDecay` method, whose specifications are *exclusive* and forbids only process trees whose set of vertices P equals $\{V_1, V_2, \dots, V_n\}$.

The **Hammer** library allows the user to specify multiple form factor ‘schemes’ to be used in reweighting. A form factor scheme is a set of FF parameterization

choices for each hadronic transition involving form factors, and is labelled by a ‘scheme name’. These schemes are set by the method `Hammer::addFFScheme`, which takes a scheme name plus a map from hadronic string representation to FF parametrization. The hadronic string follows the same syntax and uses the same particle symbols as for vertex strings. For example,

```
ham.addFFScheme("Scheme1", {{"BD", "BLPR"},
                             {"BD*", "BLPR"}});
ham.addFFScheme("Scheme2", {{"BD", "BGL"}, {"BD*",
                             "CLN"}});
```

declares two different FF schemes, choosing BLPR for both $B \rightarrow D$ and $B \rightarrow D^*$ form factors in `"Scheme1"`, and a mixture of schemes for `"Scheme2"`. Separate histograms and event weights are generated for each scheme name. The list of available FF parametrizations are provided in Table 3. The hadronic strings are charge sensitive, hence, e.g., $\{{"B+D"}, {"BLPR"}\}$ versus $\{{"BOD"}, {"CLN"}\}$ assigns two different FF parametrizations to charged and neutral $B \rightarrow D$ decays. Specification of the form factor schemes used to generate the MC sample, i.e., the denominator or input form factors, must be specified in order for **Hammer** to be able to generate the reweighting tensors. These schemes are specified by the method `Hammer::setFFInputScheme`.

Units of the input MC sample may/should be specified via `Hammer::setUnits`, for instance `ham.setUnits("MeV")`. The default is GeV.

The **Hammer** library permits the user to declare particular vertices, in either the denominator or numerator amplitude, to be evaluated as pure phase space. This is achieved by the method `Hammer::addPurePSVertices`, which takes a set of string vertices as an argument, and an optional enum `WTerm` taking values `COMMON` (default), `NUMERATOR`, or `DENOMINATOR`. As an example

```
ham.addPurePSVertices({"TauMuNuNu", "D*+DPi"});
ham.addPurePSVertices({"D*DGamma"},
                       WTerm::DENOMINATOR);
```

requests all $\tau \rightarrow \mu\nu\nu$ and $D^{*+} \rightarrow D\pi$ vertices in the numerator and all $D^* \rightarrow D\gamma$ vertices in the denominator, to be evaluated as phase space. How these requests are enforced is subject to detailed rules explained in the manual. The library employs the pure phase space definition

$$\frac{1}{\prod_k |\{s_k\}|} \sum_{s_i, r_j} |\mathcal{M}_{s_1, \dots, s_n; r_1, \dots, r_m}|^2 = 1 \times (m^{6-2n}), \quad (\text{A.1})$$

where s_i (r_i) are incoming (outgoing) quantum numbers, $|\{s_k\}|$ is the number of states of s_k , m is the mass of the parent particle in the vertex, and n the number of daughters.

Once all specifications are declared (include histograms, as below), containers are initialized by `Hammer::initRun()`. After invocation of `ham.initRun()`, manipulation of the FF default settings may be achieved via `setOptions`, which takes YAML⁵ format arguments. For instance,

```
ham.setOptions("BtoDBGL: {ChiTmB2: 0.01, ChiL: 0.002}");
```

changes the two BGL outer function parameters from their default settings. (Note that the YAML key for the relevant FF class has a "to" inserted in the hadronic transition, e.g., "BtoDBGL", rather than "BDBGL", to make it clear we are identifying settings for a particular class – the $B \rightarrow D$ BGL class – and not a process.)

By default the library computes the total rate (or looks up a partial width) the first time each unique vertex is encountered in a run. This behavior may be disabled, e.g., if the required integration is multidimensional and time consuming, via `ham.setOptions("ProcessCalc: {Rates: false}").`

To permit full flexibility in FF settings, duplication of the same FF class is permitted and may be invoked by adding a token to a FF parametrization name in `addFFScheme`, separated by an underscore. For instance, one may declare

```
ham.addFFScheme("Scheme1", {{"B+D", "BGL_1"}, {"B0D", "BGL_2"}, ...});
ham.addFFScheme("Scheme2", {{"BD", "BGL_3"}, ...});
```

This example allows independent manipulation of the BGL parameterization for each of the charged versus neutral modes in the same scheme, or between different schemes: After `initRun`, a succeeding `ham.setOptions("BtoDBGL_2:...")` would affect only the neutral B parametrization in "Scheme1".

Various other additional specification features are provided by the library, including e.g. specialization of Wilson coefficients to a particular global choice in the event tensor weight calculations. Specifications may also be declared through a card interface, as shown in `demo...card.cc` example programs provided with the source code.

A.3: Histogramming

Histograms of arbitrary dimensionality may be created by the `Hammer` library. In general, histogram bins contain event weight tensors (or direct products of them if there multiple processes in the event).

⁵See yaml.org.

A histogram is declared by `Hammer::addHistogram`, which takes as arguments a name string and either: a vector of dimensions, a bool for under/overflow and a vector of ranges; or a vector of bin edges and a bool for under/overflow. The method `addHistogram` does not create a single histogram, but rather a *histogram set*: A separate histogram is created for each unique event cascade and in turn for each FF scheme name specified by `addFFScheme`. For instance

```
ham.addHistogram("q2VsEmu", {20, 15}, false, {{3., 12.}, {0, 2.5}});
ham.addHistogram("q2VsEmu", {{3., 5., 9., 12.}, {0, 1, 2.5}}, true);
```

The first declaration creates a *histogram set* each with 20×15 bins, no under/overflow, binned uniformly over the respective ranges 3–12 and 0–2.5 (in appropriate units). With reference to the above `addFFScheme` example, this histogram set contains one histogram for each combination of either "Scheme1" or "Scheme2" with each unique $B \rightarrow D$ decay cascade. The second declaration similarly creates a set of 3×2 histograms with non-uniform bins and additional under/overflow bins.

Filling of histograms for a specific event is performed by `Hammer::fillEventHistogram`, which takes the histogram name and the values of the observables corresponding to each histogram dimension. For example, `ham.fillEventHistogram("q2VsEmu", {4., 0.5})` fills the appropriate bin element for the "q2VsEmu" histograms belonging to the event being processed, and fills the relevant histograms for each declared FF scheme name. (Invocations of `fillEventHistogram` must occur before `Hammer::processEvent`, discussed in Sec. A.4 below.) If `fillEventHistogram` is not invoked for a particular histogram for a particular event, the events tensor weight is not added to the histogram. When the under/overflow bool is set to `false`, events outside the bin ranges are ignored by `fillEventHistogram`.

Computation of the weight-squared uncertainties is off by default. This may be enabled globally via the options setting `ham.setOptions("Histos: {KeepErrors: true}").` However, for computational speed and/or memory efficiency, it may be instead enabled or disabled for individual histograms via `Hammer::keepErrorsInHistogram`, which takes the name of the histogram as an argument, and a bool. For instance

```
ham.keepErrorsInHistogram("q2VsEmu", true);
```

enables weight-squared computation for this particular histogram. This method should be invoked before `initRun`.

Various additional histogramming methods are provided by the library, that enable histogram compression, projection, and Wilson coefficient or FF special-

ization. These permit reduction of memory requirements or speed enhancements, and are detailed in the manual.

A.4: Processing

An event may contain multiple instances of `Process`, in order to account for the fact that a single event may feature, e.g., two B decay processes. The `Event` class is initialized by `Hammer::initEvent()`, which may take an optional initial event weight (this can also be set by `Hammer::setEventBaseWeight()`). `Process` instances are added by `Hammer::addProcess(proc)` which also returns a hash ID of the process. If the process is not allowed according to the `includeDecay` or `forbidDecay` specifications, the returned hash ID is zero, and the process is not added to the relevant `Event` containers.

Once a process is added, it is automatically initialized, which chiefly involves: calculating the signatures of each vertex in the decay cascade; identifying the various subamplitudes making up the cascade, as well as relevant form factor parametrizations and vertex decay rates, for both the numerator/output and denominator/input; and calculating the total rate for the vertex (this is done only the first time each unique vertex is encountered, i.e., only once per run per unique vertex and per FF scheme). The amplitude tensors and form factors are not computed, however, until the invocation of `Hammer::processEvent`.

Once all processes are added and relevant histograms (if any) have been denoted to be filled, the weights are actually computed and added to the histogram (if any) bins by invocation of `processEvent`. A pseudo-example on a single event with a set of processes might look like

```
ham.initEvent();
bool isAllowed = false;
//Create a set of Process, via addParticle and
//addVertex
for(Process& proc: processes){
    auto procID = ham.addProcess(proc)
    if(procID != 0){
        //Calculate observables, fill histograms
        isAllowed = true;
    }
}
if(isAllowed){ ham.processEvent(); }
```

which might be emplaced in a larger loop over a set of events.

A.5: Setting Wilson coefficients and form factors

Once `processEvent` is completed, the event weight may be retrieved by `Hammer::getWeight`, that takes the FF scheme name. For instance `ham.getWeight("Scheme1")` computes the *currently loaded event weight* for the *currently specified WCs and FFs*. The latter may be set via `Hammer::setWilsonCoefficients` and `setFFEigenvalues`.

The method `setWilsonCoefficients` takes a string that identifies which operator WCs are being set, and either a vector of the WC values or a map. The default WC settings are the SM. A typical example of the usage of this method is

```
ham.setWilsonCoefficients("BtoCTauNu",
    {{ "S_qLlL", 1.}, {"T_qLlL", 0.25}});
```

where the first argument specifies $b \rightarrow c\tau\nu$ four-Fermi WCs are being set, and the second argument is a `std::map<std::string, std::complex<double>>` of each WC to its desired value. The full list of WCs and their definitions is supplied in the manual. An optional third argument is the `WTerm` enum, that declares whether the evaluation should be applied to the numerator and/or denominator (numerator by default). As an alternative, one may instead pass as second argument a `std::vector<std::complex<double>>`, corresponding to the ordered basis

```
{"SM", "S_qLlL", "S_qRlL", "V_qLlL", "V_qRlL",
 "T_qLlL", "S_qLlR", "S_qRlR", "V_qLlR",
 "V_qRlR", "T_qRlR"};
```

with the conventions for these WCs shown in Table 1. It is important to note that the `setWilsonCoefficients` method, when taking a `std::map`, produces *incremental* settings changes. A subsequent invocation `ham.setWilsonCoefficients("BtoCTauNu", {{ "S_qLlL", 0.5}})` will result in $S_{qLlL} = 0.5$ and $T_{qLlL} = 0.25$. The method `resetWilsonCoefficients` resets the corresponding WCs to the default SM values.

The FF eigenvectors are (re)set via the method `Hammer::setFFEigenvalues` (`resetWilsonCoefficients`) in a similar way, identifying the FF eigenvectors to be set via the FF class prefix such as `"BtoD"` and the parametrization name. A typical example of the usage of this method is

```
ham.setFFEigenvalues("BtoD*", "BGLVar",
    {{ "delta_a1", 0.1}, {"delta_b1", -0.05}});
```

See the manual for definitions of currently implemented FF variational classes (typically denoted with a suffix `"Var"`).

A.6: Retrieval

Once all events or histograms have been processed (or reloaded from a file, see Sec. A.9), the user

may retrieve a specific histogram with the method `Hammer::getHistogram`, that takes a histogram name and a FF scheme name. NP choices must be specified first via `setWilsonCoefficients`, as must FF uncertainties via `setFFEigenvectors`. For example,

```
auto histo = ham.getHistogram("q2VsEmu", "Scheme2");
```

would contract the bin weights with the specified NP Wilson coefficients (and FF eigenvectors, if any) for each histogram in the "q2VsEmu" histogram set with FF scheme "Scheme2", and then combines them together into a single final histogram. This contracted histogram output `histo` is a (row-major) flattened vector of `BinContents` structs. This struct has members `sumWi`, `sumWi2` and `n` for sum of weights, sum of squared weights and number of events in the bin, respectively. (By contrast, the method `getHistograms` extracts all histograms of a specific name and scheme, producing a map of event hash IDs to histogram for all available "q2VsEmu" histograms with the FF scheme "Scheme2".)

Integrated rates or partial widths for a specific vertex may be retrieved via `Hammer::getRate`. The vertex is specified via either a vertex string, or the parent and daughter PDG codes, plus an FF scheme. (Partial widths are returned in the units specified by `Hammer::setUnits`; the default is GeV.) For example

```
ham.getRate(511, {-413, -14, 13}, "Scheme2");
ham.getRate("B0*-MuNu", "Scheme2");
```

both return the partial width for the $B^0 \rightarrow D^{*-}\mu^+\nu$ vertex, using the form factor parameterization specified in "Scheme2", and whatever WCs or FF uncertainties have been specified. (The `getRate` method is charge conjugate sensitive, so the vertex string must specify sufficient charges to make the vertex charge unique. For example, writing just "B0*MuNu" would correspond to not only $B^0 \rightarrow D^{*-}\mu^+\nu$, but also the very heavily suppressed $B^0 \rightarrow D^+\mu^-\bar{\nu}$.) The method `getDenominatorRate` similarly returns the partial width according to the specified denominator/input FF parametrization chosen in `setFFInputScheme`, and the denominator/input WCs or FF eigenvectors.

A.7: Multithreading

The library has the ability to perform lock-free parallelization of the `getHistogram(s)` and `getWeight` evaluations. This requires use of the thread local methods `setWilsonCoefficientsLocal` and `setFFEigenvectorsLocal` to set the desired WC or FF uncertainties. These ...Local methods take the same syntax as `setWilsonCoefficients` and `setFFEigenvectors`, but with different behaviour:

They do not set the values incrementally from the current settings, but always increment from the SM and zero FF uncertainties, respectively. Global values of the WCs or FF variations are unaffected by the ...Local methods.

A.8: Saving

`Hammer` provides the ability to store header settings, generated event weights, histograms, and/or rates in binary buffers for later retrieval and reprocessing. These buffers are built on the cross-platform serialization library `flatbuffers`⁶: The buffer structs `Hammer::IOBuffer` and `Hammer::RootIOBuffer` permit writing/reading of `Hammer` internal objects using C++ binary files and ROOT trees, respectively.

In order to save a buffer, an `ofstream` outfile must first be designated. For example, `ofstream outFile("./DemoSave.dat", ios::binary)`. The methods `Hammer::saveRunHeader`, `saveEventWeights`, `saveRates`, `saveHistogram` may be used to save: specification settings (like `includeDecay` etc; the process weight(s) of the event currently loaded in memory (this should be invoked inside an event loop, after `processEvent`); the computed rates; and histograms. Each of these methods returns a `IOBuffer`, which can be stored as sequential records in the buffer via an ostream operator. For example,

```
outFile << ham.saveRunHeader();
outFile << ham.saveHistogram("q2VsEmu");
```

writes the declared run header, with all its settings, into an `IOBuffer` and passes it as a record into the buffer, and then does the same for the histogram "q2VsEmu". The record types are labelled by an `char` enum `Hammer::RecordType` with values `UNDEFINED = 'u'`, `HEADER = 'b'`, `EVENT = 'e'`, `HISTOGRAM = 'h'`, `HISTOGRAM_DEFINITION = 'd'`, and `RATE = 'r'`. A histogram is always saved sequentially as a definition record then the histogram data record. The `saveHistogram` method may optionally take additional arguments – such as an FF scheme name – in order to save only part of an entire histogram set; see the manual for further details.

Saving a buffer in ROOT format is achieved by passing the `IOBuffer` output of the `save...` methods into a `RootIOBuffer`, that may then be stored in a ROOT `TTree`. Explicit implementations of this functionality are provided in various `demo...root.cc` example programs.

⁶google.github.io/flatbuffers.

A.9: Reloading and merging

Buffer records may be loaded from a declared `ifstream` infile into an `IOBuffer` via an `istream` operator. For example,

```
ifstream inFile("./DemoSave.dat", ios::binary);
Hammer::IOBuffer
    buf{Hammer::RecordType::UNDEFINED, 0ul,
        nullptr};
inFile >> buf;
ham.loadRunHeader(buf);
```

attempts to load the first buffer record as a run header (returning `false` if this record is of a different type).

It is the responsibility of the user to curate the logic and order under which a buffer is saved and then read. For example, if a block of histograms have been saved before a set of rate records, then

```
while(buf.kind != Hammer::RecordType::RATE) {
    if(buf.kind == Hammer::RecordType::HISTOGRAM) {
        ham.loadHistogram(buf);
    }
    if(buf.kind ==
        Hammer::RecordType::HISTOGRAM_DEFINITION){
        ham.loadHistogramDefinition(buf);
    }
    inFile >> buf;
}
```

would read through the buffer, with the method `Hammer::loadHistogram` loading all the histograms, and `Hammer::loadHistogramDefinition` all the histogram definitions, that are found before reaching the block of saved rates. The method `loadRates` behaves similarly to `loadHistogram`.

Once an object is loaded, it behaves just as the originally computed instance. So one may invoke `getHistogram` for a reloaded histogram as described in Sec. A.5.

Event weights can be reloaded via `loadEventWeights`. This permits recreating the original event loop provided `initEvent` and `processEvent` are called appropriately. For example, on a block of saved event records

```
while(buf.kind == Hammer::RecordType::EVENT) {
    ham.initEvent();
    ham.loadEventWeights(buf);
    double q2 = ...; //Calculate q^2 from known
        kinematic event information
    ham.fillEventHistogram("Q2", {q2});
    ham.processEvent();
    inFile >> buf;
}
```

would permit reprocessing of saved event weights into a newly created "Q2" histogram.

Loading a buffer in ROOT format is achieved by reading the `RootIOBuffer` stored in a `TTree` into an `IOBuffer` that can be passed to the `load...` methods. Explicit

implementations of this functionality are provided in various `demo...root.cc` example programs.

In order to permit parallelization of initialization runs, the `load...` methods accept an additional bool, to specify whether to *merge* the buffer contents with existing objects in memory (`true`), or overwrite them (`false`, default). Merging of histograms occurs if two histograms are loaded with a matching name. This merging is *additive* for histograms in each histogram set with the same FF scheme and hash IDs, and otherwise results in the new unique histograms being *appended* to the existing histogram set. (If one wishes instead to overwrite a histogram one may instead first invoke `removeHistogram`, and then reload the desired components of the histogram set.)

The methods `loadEventWeights` and `loadRates` behave similarly. For weights (rates) with matching hash IDs, merging permits appending of process weights (rates) computed with new form factor schemes to the process weights (rates). Finally, `loadRunHeader` permits merging of two sets of header specifications into their union. More details are provided in the manual.

References

1. P. Krawczyk and S. Pokorski, *Strongly Coupled Charged Scalar in B and t Decays*, *Phys. Rev. Lett.* **60** (1988) 182.
2. P. Heiliger and L. M. Sehgal, *Semileptonic Decays of B Mesons into $\tau\nu_\tau$* , *Phys. Lett.* **B229** (1989) 409–417.
3. J. Kalinowski, *Semileptonic Decays of B Mesons into $\tau\nu_\tau$ in a Two Higgs Doublet Model*, *Phys. Lett.* **B245** (1990) 201–206.
4. B. Grzadkowski and W.-S. Hou, *Solutions to the B meson semileptonic branching ratio puzzle within two Higgs doublet models*, *Phys. Lett.* **B272** (1991) 383–390.
5. Y. Grossman and Z. Ligeti, *The Inclusive $\bar{B} \rightarrow \tau\bar{\nu}X$ decay in two Higgs doublet models*, *Phys. Lett.* **B332** (1994) 373–380, [[hep-ph/9403376](#)].
6. M. Tanaka, *Charged Higgs effects on exclusive semitauonic B decays*, *Z. Phys.* **C67** (1995) 321–326, [[hep-ph/9411405](#)].
7. W. D. Goldberger, *Semileptonic B decays as a probe of new physics*, [[hep-ph/9902311](#)].
8. ALEPH Collaboration: D. Buskulic *et al.*, *Measurement of the $b \rightarrow \tau^-\bar{\nu}X$ branching ratio*, *Phys. Lett.* **B298** (1993) 479–491.
9. HFLAV Collaboration: Y. S. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ -lepton properties as of 2018*, [[arXiv:1909.12524](#)].

10. BaBar Collaboration: J. P. Lees *et al.*, *Measurement of an Excess of $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, *Phys. Rev.* **D88** (2013) 072012, [[arXiv:1303.0571](#)].
11. LHCb Collaboration: R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)$* , *Phys. Rev. Lett.* **115** (2015) 111803, [[arXiv:1506.08614](#)]. [Addendum: *Phys. Rev. Lett.* 115, no.15, 159901 (2015)].
12. Belle Collaboration: M. Huschle *et al.*, *Measurement of the branching ratio of $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, *Phys. Rev.* **D92** (2015) 072014, [[arXiv:1507.03233](#)].
13. Z. Ligeti, M. Papucci, and D. J. Robinson, *New Physics in the Visible Final States of $B \rightarrow D^{(*)}\tau\nu$* , *JHEP* **01** (2017) 083, [[arXiv:1610.02045](#)].
14. S. Duell, F. Bernlochner, Z. Ligeti, M. Papucci, and D. Robinson, *HAMMER: Reweighting tool for simulated data samples*, *PoS ICHEP2016* (2017) 1074.
15. D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth.* **A462** (2001) 152–155.
16. F. U. Bernlochner, Z. Ligeti, M. Papucci, and D. J. Robinson, *Combined analysis of semileptonic B decays to D and D^* : $R(D^{(*)})$, $|V_{cb}|$, and new physics*, *Phys. Rev.* **D95** (2017) 115008, [[arXiv:1703.05330](#)].
17. N. Isgur and M. B. Wise, *Weak Decays of Heavy Mesons in the Static Quark Approximation*, *Phys. Lett.* **B232** (1989) 113–117.
18. N. Isgur and M. B. Wise, *Weak Transition Form-factors Between Heavy Mesons*, *Phys. Lett.* **B237** (1990) 527–530.
19. E. Eichten and B. R. Hill, *An Effective Field Theory for the Calculation of Matrix Elements Involving Heavy Quarks*, *Phys. Lett.* **B234** (1990) 511–516.
20. H. Georgi, *An Effective Field Theory for Heavy Quarks at Low-energies*, *Phys. Lett.* **B240** (1990) 447–450.
21. BaBar Collaboration: J. P. Lees *et al.*, *Evidence for an excess of $\bar{B} \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ decays*, *Phys. Rev. Lett.* **109** (2012) 101802, [[arXiv:1205.5442](#)].
22. LHCb Collaboration: R. Aaij *et al.*, *Measurement of the CKM angle γ from a combination of LHCb results*, *JHEP* **12** (2016) 087, [[arXiv:1611.03076](#)].
23. G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J.* **C71** (2011) 1554, [[arXiv:1007.1727](#)]. [Erratum: *Eur. Phys. J.* C73,2501(2013)].
24. I. Doršner, S. Fajfer, A. Greljo, J. F. Kamenik, and N. Košnik, *Physics of leptoquarks in precision experiments and at particle colliders*, *Phys. Rept.* **641** (2016) 1–68, [[arXiv:1603.04993](#)].
25. M. Freytsis, Z. Ligeti, and J. T. Ruderman, *Flavor models for $\bar{B} \rightarrow D^{(*)}\tau\bar{\nu}$* , *Phys. Rev.* **D92** (2015) 054018, [[arXiv:1506.08896](#)].
26. J. Aebischer, J. Kumar, P. Stangl, and D. M. Straub, *A Global Likelihood for Precision Constraints and Flavour Anomalies*, *Eur. Phys. J.* **C79** (2019) 509, [[arXiv:1810.07698](#)].
27. P. Asadi and D. Shih, *Maximizing the Impact of New Physics in $b \rightarrow c\tau\nu$ Anomalies*, *Phys. Rev.* **D100** (2019) 115013, [[arXiv:1905.03311](#)].
28. D. Bardhan and D. Ghosh, *B -meson charged current anomalies: The post-Moriond 2019 status*, *Phys. Rev.* **D100** (2019) 011701, [[arXiv:1904.10432](#)].
29. S. Bhattacharya, S. Nandi, and S. Kumar Patra, *$b \rightarrow c\tau\nu$ Decays: a catalogue to compare, constrain, and correlate new physics effects*, *Eur. Phys. J.* **C79** (2019) 268, [[arXiv:1805.08222](#)].
30. C. Murgui, A. Peñuelas, M. Jung, and A. Pich, *Global fit to $b \rightarrow c\tau\nu$ transitions*, *JHEP* **09** (2019) 103, [[arXiv:1904.09311](#)].
31. A. Azatov, D. Bardhan, D. Ghosh, F. Sgarlata, and E. Venturini, *Anatomy of $b \rightarrow c\tau\nu$ anomalies*, *JHEP* **11** (2018) 187, [[arXiv:1805.03209](#)].
32. D. Buttazzo, A. Greljo, G. Isidori, and D. Marzocca, *B -physics anomalies: a guide to combined explanations*, *JHEP* **11** (2017) 044, [[arXiv:1706.07808](#)].
33. F. Bernlochner, S. Duell, Z. Ligeti, M. Papucci, and D. J. Robinson, “An introduction to hammer: Helicity amplitude module for matrix element reweighting.” [hammer.physics.lbl.gov](#).
34. D. Scora and N. Isgur, *Semileptonic meson decays in the quark model: An update*, *Phys. Rev.* **D52** (1995) 2783–2812, [[hep-ph/9503486](#)].
35. N. Isgur, D. Scora, B. Grinstein, and M. B. Wise, *Semileptonic B and D Decays in the Quark Model*, *Phys. Rev.* **D39** (1989) 799–818.
36. B. Grinstein and A. Kobach, *Model-Independent Extraction of $|V_{cb}|$ from $\bar{B} \rightarrow D^*\ell\bar{\nu}$* , *Phys. Lett.* **B771** (2017) 359–364, [[arXiv:1703.08170](#)].
37. C. G. Boyd, B. Grinstein, and R. F. Lebed, *Model independent determinations of $\bar{B} \rightarrow D\ell\bar{\nu}$, $D^*\ell\bar{\nu}$ form-factors*, *Nucl. Phys.* **B461** (1996) 493–511, [[hep-ph/9508211](#)].
38. C. G. Boyd, B. Grinstein, and R. F. Lebed, *Precision corrections to dispersive bounds on*

- form-factors, *Phys. Rev.* **D56** (1997) 6895–6911, [[hep-ph/9705252](#)].
39. I. Caprini, L. Lellouch, and M. Neubert, *Dispersive bounds on the shape of $\bar{B} \rightarrow D^{(*)} \ell \bar{\nu}$ form-factors*, *Nucl. Phys.* **B530** (1998) 153–181, [[hep-ph/9712417](#)].
 40. J. H. Kuhn and E. Mirkes, *Structure functions in tau decays*, *Z. Phys.* **C56** (1992) 661–672. [Erratum: *Z. Phys.* C67,364(1995)].
 41. O. Shekhovtsova, T. Przedzinski, P. Roig, and Z. Was, *Resonance chiral Lagrangian currents and τ decay Monte Carlo*, *Phys. Rev.* **D86** (2012) 113008, [[arXiv:1203.3955](#)].
 42. I. M. Nugent, T. Przedzinski, P. Roig, O. Shekhovtsova, and Z. Was, *Resonance chiral Lagrangian currents and experimental data for $\tau^- \rightarrow \pi^- \pi^- \pi^+ \nu_\tau$* , *Phys. Rev.* **D88** (2013) 093012, [[arXiv:1310.1053](#)].
 43. A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, *Semileptonic B decays to excited charmed mesons*, *Phys. Rev.* **D57** (1998) 308–330, [[hep-ph/9705467](#)].
 44. A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, *Model independent results for $B \rightarrow D_1(2420) \ell \bar{\nu}$ and $B \rightarrow D_2^*(2460) \ell \bar{\nu}$ at order $\Lambda_{\text{QCD}}/m_{c,b}$* , *Phys. Rev. Lett.* **78** (1997) 3995–3998, [[hep-ph/9703213](#)].
 45. F. U. Bernlochner, Z. Ligeti, and D. J. Robinson, *Model independent analysis of semileptonic B decays to D^{**} for arbitrary new physics*, *Phys. Rev.* **D97** (2018) 075011, [[arXiv:1711.03110](#)].
 46. F. U. Bernlochner and Z. Ligeti, *Semileptonic $B_{(s)}$ decays to excited charmed mesons with e, μ, τ and searching for new physics with $R(D^{**})$* , *Phys. Rev.* **D95** (2017) 014022, [[arXiv:1606.09300](#)].
 47. M. Pervin, W. Roberts, and S. Capstick, *Semileptonic decays of heavy lambda baryons in a quark model*, *Phys. Rev.* **C72** (2005) 035201, [[nucl-th/0503030](#)].
 48. F. U. Bernlochner, Z. Ligeti, D. J. Robinson, and W. L. Sutcliffe, *Precise predictions for $\Lambda_b \rightarrow \Lambda_c$ semileptonic decays*, *Phys. Rev.* **D99** (2019) 055008, [[arXiv:1812.07593](#)].
 49. F. U. Bernlochner, Z. Ligeti, D. J. Robinson, and W. L. Sutcliffe, *New predictions for $\Lambda_b \rightarrow \Lambda_c$ semileptonic decays and tests of heavy quark symmetry*, *Phys. Rev. Lett.* **121** (2018) 202001, [[arXiv:1808.09464](#)].